PERSONAL RESUME

基本信息及教育背景

个人主页地址:https://mahonghao1.github.io/



姓名：麻洪豪　　政治面貌：中共党员

手机号码：18177188225　　邮箱:Mahonghao11@outlook.com

求职意向：具身智能数据处理 /数据pipeline / 数据平台开发

2023.09-2027.06　　上海应用技术大学　　机器人工程专业

GPA：3.8/4（专业前5%）|主修课程：机器人控制技术、传感器与检测技术、移动机器人技术、机电一体化系统设计

实习经历

具身数据算法工程师　　擎仓机器人(Quic Robot)　　2026.10–至今

- 真机数采部署：面向真机机械臂工作站，实现 QRDF/LeRobot 双格式同步双写与三相机适配；针对长时采集深度流内存溢出，设计分段有界缓冲与原子落盘方案，单条内存峰值 25.9GiB 降至 63MiB,支撑 1510 秒长时序采集。
- 数据对账闭环：针对上传链路多阶段流转、缺陷数据难定位问题，基于 SQLite 作业状态库实现 2,3000+ 作业分阶段对账与根因分类，以 SHA-256 完整性校验处置 53 条不一致数据，沉淀“采集-上传-准入”可回溯质检机制。
- 平台算子开发：设计暂存自动回收算子，以限额分批、先删后标记、失败自愈策略实现数据零误删，610 条作业生产规模仿真 20/20 通过、预计回收 400TB+;修复上传队列 3 处异常缺陷，回归 678 项测试零新增失败。
- 采集界面优化：排查采集端操作无响应类问题，定位为断连静默失败与回执无超时机制所致，实施断连分级、超时兜底、相机保帧等 10 项改进，前端 60 项测试全部通过。
- 客户现场联调：参与华润 WMS 与擎仓 WCS 四接口联调与 SM2 国密双向验签，真机任务闭环跑通；定位 38 条任务下发全失败根因为商品档案尺寸/重量字段缺失，输出修复方案闭环。

数据交付工程师实习生　　光轮智能（Lightwheel）　　2026.6–2026.10

- 数据格式转换：承接 Wuji 原始采集数据，独立开发 LeRobot v3.0 转换与校验工具，对齐位姿、相机参数、MANO 手部参数等多模态字段，输出 3323 帧标准数据集并完成 schema/字段一致性校验。
- 交付脚本开发：面向 Sample、Plum、Papaya、Jackfruit 等客户交付场景，迭代 Python 数据导出脚本，自动组装 task.zip/fixed_task.zip、metadata、pose、点云、相机视频等标准交付产物。
- 链路稳定性优化：补齐 workflow_run_id、batch_version、operator 等元数据兜底逻辑，修复平台用户查询失败导致的导出中断；参与 Docker 镜像构建、跨仓库推送、容器自测与真实交付包校验。

具身智能数据工程实习生　　中国电信人工智能研究院(Tele ai)　　2026.1-2026.6

- 管线搭建落地：面向 VLA 具身预训练需求，基于 Python 搭建自动化多模态数据处理流水线，制定 8 类复杂场景采集规范，处理 PB 级超 4 亿帧样本，数据交付合格率稳定 98% 以上。
- 数据质检体系:搭建全覆盖自动化校验规则，检测时间戳同步、缺帧重复、异常轨迹等脏数据，减少人工复核成本，为强化学习真机实验提供高质量训练素材。
- 项目成果支撑:参与 VLA 模型真机推理与 RL 策略验证，输出标准化实验数据集，助力团队开源项目与 Arxiv 学术论文顺利发布。论文：https://arxiv.org/abs/2604.27472 | 代码：https://github.com/TeleHuman/PRTS

智能车舱数据工程实习生　　商汤科技(Sensetime)　　2025.7-2026.1

- 长尾样本扩充：针对车载 DMS/OMS 模型误检问题，设计多模态采集协议，累计产出 2TB 标注数据，补充 500GB 稀缺长尾样本，使模型量产误检率降低 10%。
- 采集流程自动化：开发 Python 自动化采集脚本，实现设备启停、断点续传、任务自动归档，单次采集人工操作时长缩减 40%，脚本复用至 3 款新车项目。
- 多传感时序对齐：设计多路相机、麦克风同步方案，输出时序对齐多模态数据流，为视线跟踪、手势识别模型提供规整训练输入。

项目/专利经历

RoboData Workbench｜具身智能数据生产线

github.com/MAHONGHAO1/Embodied-AI-Data-Processing-Platform

- 全流水线搭建：独立搭建机器人演示数据处理流水线：接入→清洗筛选→标注审核→质检→ 转换→打包交付，人工与自动节点混合编排；覆盖仿真与真机两类来源，真机数据全量 50 条／32,068 帧接入并跑通闭环，HDF5 批次进来、官方训练就绪数据集与交付包出去，全程平台内完成。
- 质量门禁前置：将交付质检经验沉淀为流水线内建规则——时间戳同步、缺帧重复、异常轨迹、NaN 行级定位；执行失败轨迹与数据缺陷分开处置，失败轨迹保留为训练负样本资产而非丢弃；质检未通过的批次被强制拦截，不进转换、不进交付。
- 交付可信验证：结果经官方 LeRobot 库离线加载＋PyTorch DataLoader 批读取复核；交付包独立解包+SHA-256 防篡改

AI 智选基金助手|个人投资辅助决策系统

Github.com/MAHONGHAO1/AI-Fund-Management-Platform

- 多源数据容错：集成 AkShare 与东方财富双数据源，针对公开接口限频、字段不稳定问题，设计自动重试、跨源切换、模拟数据兜底的多层容错链路，保障看板在接口异常时数据持续可用。
- 量化评分模型：参考晨星评价体系实现五维加权评分(收益 30%、风控 25%、风险调整收益 25%、稳定性 10%、相对排名 10%),输出 0–10 分与星级评级，支持基金横向可比与收益-回撤散点、五维雷达可视化。
- AI 建议生成：接入通义千问生成结构化投资分析(趋势判断、操作建议、风险提示)，设计 API 故障自动降级本地规则引擎的兜底机制，支持定投回测模拟与 Markdown/HTML 报告一键导出。

实用新型专利：《一种机器人快速升降定位装置》｜第一发明人

- 提出“减速部-缓冲部-定位孔”三段式机械锁止结构，降低机器人快速升降过程中的到位偏移与冲击损伤。

竞赛与荣誉

中国国际“互联网+”大学生创新创业大赛国家级铜奖｜全国无人系统应用大赛国家级三等奖｜“西门子杯”逻辑算法赛道省部级三等奖｜知行杯省部级三等奖 | 多次获校优秀奖学金 | 先进个人 | 优秀学生干部 | 优秀学生 | 优秀团员

核心技能

- AI/具身智能： 熟悉codex/claude code/openclaw、了解基于 Transformers 的VLA模型训练与微调。
- 数据工程： 熟悉多模态数据采集、清洗、质检、标注与 LeRobot/HDF5/QRDF 格式转换，具备 PB 级机器人数据处理经验。
- 工程开发： Python、C/C++、ROS1/2、Linux、SQL、Streamlit、Plotly，具备自动化脚本开发与多传感器同步经验。
- 证书语言： 人工智能训练师、上海市信息技术水平人工智能一级、计算机 C 语言三级，英语四级，可阅读英文技术文档。

校园经历

团委志愿服务中心 **团委志愿服务中心主任** **2025.06-2026.06**

担任智能技术学部团委志愿中心主任期间，统筹多项大型活动。主导5000+人迎新活动、1000+人迎新晚会及万人招聘会的 策划执行，展现卓越的组织协调能力。负责中国国际进口博览会志愿者专项工作，组织制定招募标准并完成百余名志愿者的 精准选拔；设计并实施涵盖服务技能、应急处理等内容的系统化培训，全面提升团队专业素养；全程进行现场指导与岗位调 配，确保服务任务高质量完成，获得各方高度认可。同时，积极联动上海图书馆、上海南站等单位，搭建校地志愿服务平 台，强化社会责任担当。